Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 6812–6816, November 1984
Genetics

Directional cloning of DNA fragments at a large distance from an initial probe: A circularization method

(gene mapping/chromosome walking/suppressor tRNA)

FRANCIS S. COLLINS* AND SHERMAN M. WEISSMAN†

Departments of Human Genetics, Medicine, and Molecular Biophysics and Biochemistry, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510

Contributed by Sherman M. Weissman, July 16, 1984

ABSTRACT **The principle of a DNA cloning procedure that directionally generates genomic DNA fragments 50–2000 kilobases away from an initial probe is presented. The method depends on partial digestion of high molecular weight genomic DNA and subsequent ligation at very low concentration to generate covalent DNA circles. A library of the junction fragments from these circles can then be constructed. Biological or physical selection of the junction pieces can be achieved by incorporating a marker DNA fragment into the covalent circles. A 45-kilobase cosmid fragment has been successfully used to test the procedure. At appropriately low ligation concentrations (0.8 μg/ml), ≈90% of the ligated DNA is present as monomeric circles. Larger DNA fragments will require reducing the DNA concentration as the inverse square root of the DNA length. A suppressor tRNA gene has been tested as the selectable marker gene. Ligation of the digested circles into an amber-mutated λ phage and propagation in a sup⁻ host allows only the phage that contain junction fragments to produce plaques. Potential applications of this approach, such as mapping of complex genetic loci or moving from a linked gene toward a gene of interest, are presented and discussed.**

Techniques for cloning DNA fragments have contributed in a major way to the remarkable advances in molecular biology over the past decade. The purification of large quantities of specific fragments of DNA by these methods has allowed studies of gene sequence and function, including those of many genes whose protein products had not been previously characterized. The development of cosmid cloning techniques, which are capable of generating clones up to 45 kilobases (kb) long, extended the ability to characterize larger segments (1). Indeed, for several gene families, such as the α- and β-globin gene clusters (2, 3), the major histocompatibility complex (4), and the complement genes (5), it has been possible to generate a series of overlapping cosmid clones encompassing as much as 240 kb of contiguous DNA.

At the same time, through the techniques of somatic cell genetics (6) and *in situ* hybridization (7), it has been possible to map single-copy genes to particular human chromosomes and, with current techniques of cytogenetics, even to a small region of a particular chromosome. A significant gap still exists, however, between the resolution of these methods (≈5,000–10,000 kb) and the cloning approach.

One of the goals of molecular biology is to bridge this gap and permit the detailed mapping of relatively large areas of the genome. This is an important task in such complex genetic loci as the major histocompatibility complex, which may occupy more than 1000 kb of contiguous DNA (8), and the heavy and light chain immunoglobin loci, neither of which have yet been internally connected by standard cloning techniques (9). While mapping such large regions is theoretically possible with the technique of "chromosome walking," using overlapping λ or cosmid clones (4), in practice this procedure becomes prohibitively time-consuming when more than a few steps are required. What is needed is a method that will allow the investigator to take larger steps along the chromosome, preferably with the ability to specify in which direction the step will be taken. The purpose of this paper is to present the principle of such a method and a series of experiments, using a model system, that demonstrate its practicality.

The principle is outlined in diagrammatic form in Fig. 1. If it is desired to take steps of length *N* kb, very high molecular weight DNA is first partially digested with a frequent-cutting enzyme (such as *Mbo* I) and size-selected so that the fragments obtained are ≈*N* kb in length. The DNA is then ligated at low concentration such that ligation into circles is markedly favored over the formation of multimers. It can also be advantageous to include a marker DNA segment within the covalently linked circle, to allow identification of the junction fragments in subsequent manipulations. An example of such a marker is the amber suppressor tRNA gene (*supF*) shown in Fig. 1. The circles are digested with a second enzyme (such as *Eco*RI), generating a large number of fragments. Most of these will be contiguous DNA pieces from the genome, but a few will represent junction fragments originally arising from the two ends of an *N*-kb-long linear DNA. The selectable marker gene can then be used to generate a library of such junction fragments. In the example shown, the presence of the suppressor tRNA gene allows these fragments, when ligated into an amber-mutated bacteriophage such as λ Charon 16A (λCh16A), to be propagated on a sup⁻ host. Other λCh16A recombinant phages will not be able to grow and, therefore, it is possible to generate a library of such junction fragments and screen it with a probe representing the starting point of the desired chromosome "hop." Clones selected from the library will consist of DNA homologous to the probe separated by a suppressor tRNA gene from another DNA sequence that originally lay *N* kb away. By generating several libraries that have different values of *N*, it should be possible to move from an original location a desired distance away by screening the proper library. It is possible to influence the direction in which the hop occurs by proper selection of the relationship of the probe to genomic restriction sites (see discussion below).

MATERIALS AND METHODS

Preparation of Test DNA Fragment. To obtain a homogeneous population of large DNA fragments with no internal *Bam*HI sites, high molecular weight human genomic DNA (10) was digested to completion with the enzyme *Bam*HI and ligated to the cosmid vector pJB8 (11). After packaging and transfection, preparation of small amounts of DNA allowed identification of those cosmids that contained a single large *Bam*HI fragment and a single copy of the vector. Six cosmid

Abbreviation: kb, kilobase(s).

*Present address: Department of Medicine, University of Michigan Medical School, Ann Arbor, MI 48109.
†To whom reprint requests should be addressed.

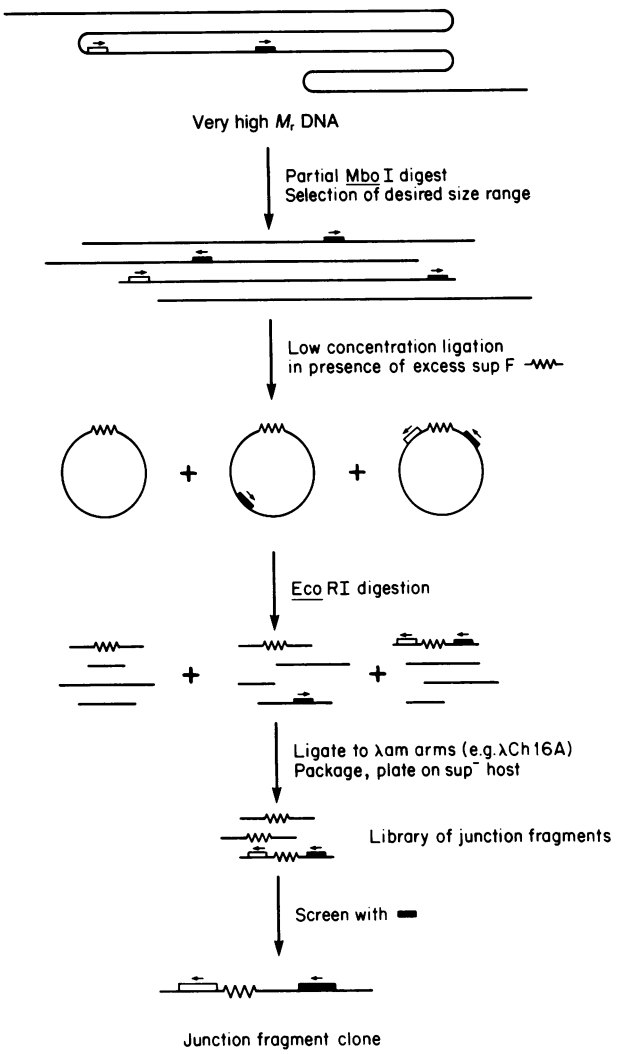

FIG. 1. Principle of the cloning procedure. The heavy bar represents the starting probe, which in the final junction-fragment clone is present along with the marker gene and another segment of DNA (open box) that was initially many kb away in the genome. In this particular example, a suppressor tRNA gene (*supF*) is used as the marker. Marker DNA segments allowing biological or physical separation of the junction pieces may also be of use. Horizontal arrows show the orientation of the junction fragment pieces relative to their original genomic arrangement.

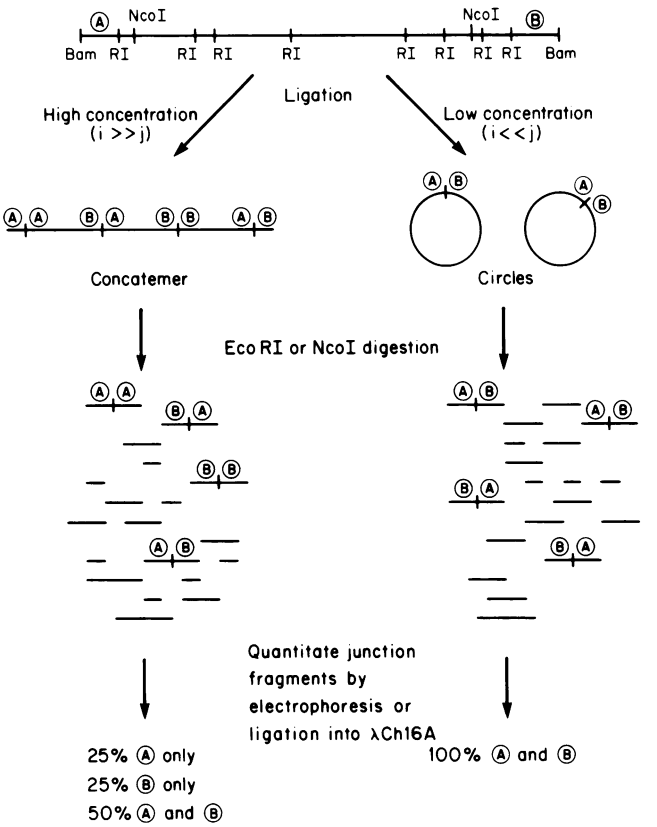

FIG. 2. Method for assessing relative efficiency of circularization. The DNA fragment at the top represents the 45-kb *Bam*HI (Bam) insert from cosmid B1-6, with *Eco*RI (RI) end pieces A and B and *Nco* I end pieces A′ and B′ (only the terminal *Nco* I sites are shown). The parameters *i* and *j* are defined in the text.

inserts were purified by *Bam*HI digestion of 100 μg of cosmid DNA and separation from the vector on a 10–40% sucrose gradient. End pieces were identified by filling in the *Bam*HI sites with the Klenow fragment of *Escherichia coli* DNA polymerase I in the presence of [α-^{32}P]dGTP (12) and preparing autoradiographs of DNA fragments electrophoresed after *Eco*RI digestion. The cosmid inserts were also digested with *Eco*RI, electrophoresed on a 0.8% agarose gel, transferred to nitrocellulose paper (13), and hybridized with nick-translated total human genomic DNA (12) to identify highly repetitive fragments. One cosmid, denoted B1-6, was found to have nonrepetitive *Bam*HI/*Eco*RI end pieces of 2.2 and 2.4 kb. These end pieces were subcloned and the 2.2- and 2.4-kb inserts, denoted probes A and B, were purified by gel electrophoresis and electroelution (12). These were found not to cross-hybridize with other fragments of B1-6.

Assay of Circularization Efficiency. The basic principle of the method is outlined in Fig. 2. One microgram of purified insert from cosmid B1-6 was ligated at concentrations ranging from 120 μg/ml to 0.6 μg/ml in 50 mM Tris·HCl, pH 7.4/10 mM $MgCl_2$/20 mM dithiothreitol/1 mM ATP containing gelatin at 50 μg/ml. Ligations were carried out at 14°C for 12–16 hr at a concentration of T4 DNA ligase (New England Biolabs) ≥1.6 units/μl. DNA was recovered by ethanol precipitation. Direct visualization of circularization efficiency was accomplished by *Nco* I digestion and gel electrophoresis; the *Nco* I end fragments A′ (3.3 kb) and B′ (6.0 kb) will, at 100% circularization efficiency, form only A′B′ (9.3 kb) but, if concatemers are present, A′A′ (6.6 kb) forms will also be seen (Fig. 3). Alternatively, the ligated DNA was digested instead with *Eco*RI; 30 ng of the digested DNA was ligated with 0.25 μg of *Eco*RI-cut, phosphatase-treated λCh16A DNA. (Boehringer Mannheim molecular biology grade calf intestinal phosphatase was used, following the method of ref. 12.) This ligated material was packaged using standard methods (3) and transfected into *supF*$^+$ strain LE392. One-third to one-half of the resulting phage plaques contained inserts derived from cosmid B1-6, with an overall efficiency of 10^7 plaques per μg of insert DNA. To determine the proportion of circles that had formed in the initial ligation, 2000 plaques were screened with probe A, stripped, and then screened with probe B. Nitrocellulose filter preparation, nick-translation of the probes, and hybridization conditions followed standard methods (12). Approximately 15% of the plaques that contained inserts included one or both of the end pieces A and B. Direct comparison of autoradiographs from sequential screenings of a given filter with probes A and B allowed ready identification of plaques that hybridized with both probes or with one only.

Purification of Suppressor tRNA Gene. The plasmid pRD69 (a gift of R. Dunn, Department of Medical Genetics, University of Toronto) consists of a synthetic tyrosine amber suppressor tRNA gene, with *Eco*RI boundary sites, cloned into

pBR322 (14). Five micrograms of pRD69 was cut with *Bam*HI, the ends of the fragments were filled in with the Klenow fragment of *E. coli* DNA polymerase I, and *Bam*HI linkers (Boehringer Mannheim; sequence G-G-G-A-T-C-C-C, so that the *Eco*RI site was destroyed) were attached (12). This DNA was ligated into pBR328 and transfected into CARD-15, which contains an amber-mutated *lacZ* gene (R. Dunn, personal communication). Plating the resulting colonies on chloramphenicol- and lactose-supplemented MacConkey agar yielded lactose-fermenting colonies containing the suppressor gene plasmid. Digestion of DNA from this plasmid, denoted pFC2, with *Bam*HI, followed by electrophoresis of the digest in 1.4% agarose and electroelution of the desired fragment (12), yielded the desired 219-base-pair *supF* gene.

Incorporation of *supF* Gene into Circles. With the concentration of the B1-6 DNA at 1.0 μg/ml, a 500-fold molar excess of purified *supF* was included in the ligation mixture prior to addition of T4 DNA ligase. Subsequent ligation, digestion with *Nco* I or *Eco*RI, and gel electrophoresis or ligation into λCh16A were as described above. Packaged material was plated on LE392, which contains its own suppressor tRNA gene, or on CA274 (15), which is sup$^-$. DNA from plaques positive for one or both of the A and B probes on LE392 was prepared, digested with *Eco*RI, electrophoresed, transferred to nitrocellulose, and hybridized with pRD69 to detect the presence of the *supF* gene in the phage inserts. Phage bearing the *supF* gene were plated on CA274 and their plaque-forming ability was compared with LE392. An alternative plating method that increased the efficiency of plaque formation of some *supF*$^+$ phage was to incubate with CA274 for 15 min at 37°C, collect the cells by centrifugation, wash them with medium, and then suspend them in a culture of LE392 just before plating.

RESULTS

A crucial step in the method we propose is the ligation of high molecular weight DNA fragments preferentially into circles rather than into concatenated multimers. Testing the efficiency of circularization required the construction of a model system as diagrammed in Fig. 2. The assay depends on analysis of the composition of junction fragments. At high DNA concentrations, the results of ligation will be a long concatemer, the junction pieces of which will be 25% AA, 50% AB, and 25% BB. At low concentrations, where the formation of circles is favored, the distribution should approach 100% AB, since this is the only possible product of a circularization. At an intermediate point,

$$\% \text{ linear} = 2\,(\%AA + \%BB) \quad [1]$$

$$\text{and } \% \text{ circular} = 100 - 2\,(\%AA + \%BB). \quad [2]$$

Direct analysis of the junction fragments is shown in Fig. 3*a*; the ligated DNA has been cut with *Nco* I. Incomplete ligation is detected by partial persistence of the A′ and B′ fragments, but the efficiency of ligation is ≥50% throughout the range of concentrations. At lower DNA ligation concentrations, the A′A′ band is fainter, indicating that more circles are formed. This decrease in the A′A′ band can also be demonstrated by Southern blotting as shown in Fig. 3*b*. At a ligation concentration of 1 μg/ml, however, the A′A′ band is not detectable, implying that circularization predominates.

More quantitative analysis of circle formation at low ligation concentrations was achieved by digesting the ligated DNA with *Eco*RI, cloning the fragments into λCh16A, sequentially screening the resulting phage with the two end pieces (A and B) of the 45-kb DNA test fragment, and applying Eq. 2 above. This assumes that AA, AB, and BB will be cloned with equal efficiencies, which is justified by their close size similarity. The results of this analysis at DNA concentrations ranging from 40 μg/ml to 0.6 μg/ml are shown in Fig. 4. No loss of ligation efficiency was noted at the lowest concentrations used. The observed efficiency of circularization is in good agreement with theoretical predictions.

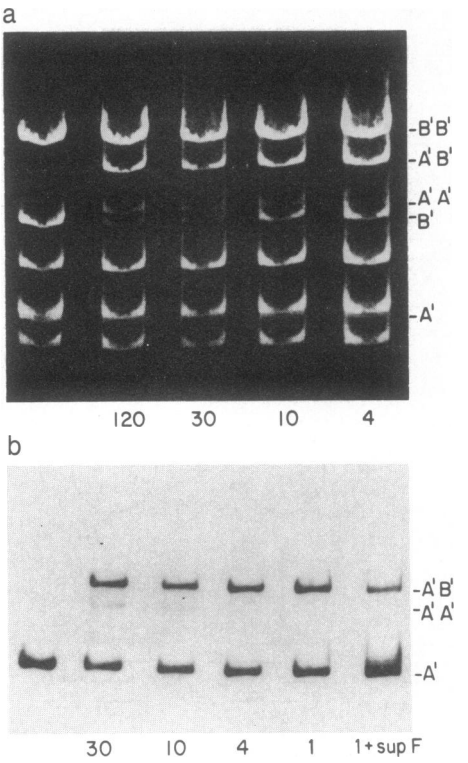

FIG. 3. Analysis of circularization efficiency by electrophoresis of *Nco* I digests of the ligated 45-kb test DNA. (*a*) Ethidium-stained gel. The left lane contained unligated DNA; samples in other lanes resulted from ligations done at DNA concentrations of 120, 30, 10, and 4 μg/ml as indicated at the bottom of the lanes. A′ and B′ are the two *Nco* I endpieces as shown in Fig. 2. At high ligation concentrations (e.g., 120 μg/ml), the molar concentrations of the A′A′ band should be 50% that of A′B′. At lower concentrations, the decrease in the A′A′/A′B′ ratio is an indication of increased circle formation. B′B′ comigrates with an internal *Nco* I fragment. (*b*) Southern blot of a similar gel using a probe for A′. Again, the A′A′ band is greatly reduced at low ligation concentrations and cannot be seen at 1 μg/ml. The sample in the right-most lane included a 500-fold molar excess of *supF* gene, added before ligation; note that this addition did not result in any reappearance of A′A′.

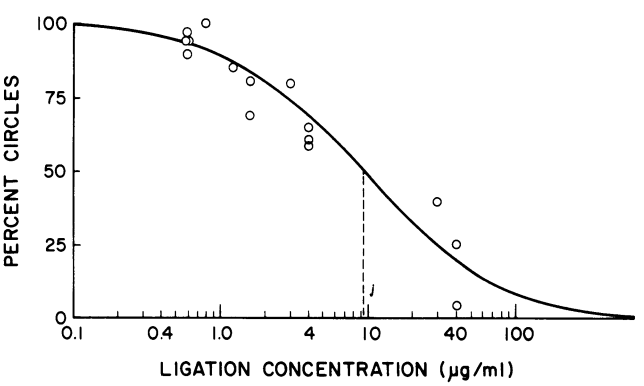

FIG. 4. Percentage of ligation into circles of a 45-kb DNA piece as a function of DNA concentration during ligation. Each point represents an independent experiment. The concentration *j* is that at which, theoretically, 50% of the ligated DNA should be circular. The solid line is based on Eq. 5 (see *Discussion*).

Incorporation of a marker was assessed by adding a 500-fold molar excess of a 219-base-pair *Bam*HI-ended tyrosine amber suppressor tRNA gene (*supF*) to the low-concentration ligation mixture. To test incorporation of *supF* into A′B′, the ligated DNA was digested with *Eco*RI and cloned into λCh16A; plating on a sup$^+$ host and screening with the A and B probes demonstrated that addition of the *supF* gene to the ligation mixture did not significantly alter either the efficiency of ligation or the percentage of the junction fragments derived from circles. DNA from minilysates of these junction-fragment plaques was digested with *Eco*RI and hybridized with *supF*. Of eight clones analyzed, six carried one to three *supF* markers separating two end fragments within the junction piece (data not shown). Unexpectedly, however, the plaque-forming ability of these *supF*-bearing phage on the sup$^-$ host CA274 was variable. In particular, the presence of the A fragment in the phage exerted a pronounced inhibitory effect on plaque formation whereas the B fragment did not have such an effect. Possibly, the nature of the sequence immediately adjacent to the 5′ end of the *supF* gene has an influence on its expression in *E. coli*. This inefficiency of plaque formation was overcome, however, by preincubating the *supF*-bearing phage with CA274; during this time the phage attach and inject their DNA. Removing the supernatant and adding LE392 as an "amplifying lawn" allowed efficient plaque formation from these phage; *supF*$^-$ phage yield <1% of their titer on LE392 by this protocol.

DISCUSSION

The success of the recombinant DNA approach in answering many important questions about the structure and function of genes has led to the need to map increasingly larger segments of the human genome. The approach diagrammed in Fig. 1 and described in the Introduction presents a method of overcoming distance limitations, by generating a clone many kb away from an initial probe without the need to characterize all of the intervening DNA. The exact distance from the initial probe can be set by the size of the partially digested high molecular weight DNA chosen for construction of the library. Size fractionation up to 60–80 kb can be done by sucrose gradient methods, but fractionating pieces larger than this may require electrophoresis at low agarose concentrations (16–18) or the pulsed-field gradient electrophoresis methods described by Schwartz, Cantor, and co-workers (19, 20). Note that DNA fragments that are unable to ligate due to damage to their ends will not confuse the subsequent analysis, since only a fragment that has *Mbo* I sites at both ends can give rise to a plaque in the final library.

The theoretical basis of ligation of a linear polymeric molecule in solution (21) has been experimentally tested for bacteriophage λ DNA (22) and smaller DNA fragments (23). Jacobson and Stockmayer (21) considered the behavior of a polymer consisting of a series of rigid segments of length b, joined by freely movable joints, and of total contour length l. The theoretical treatment results in the calculation of a parameter, j, that is the effective concentration of one end of a long polymer in the neighborhood of the other end. This parameter is given by the equation

$$j = \left(\frac{3}{2\pi lb}\right)^{3/2} \text{ ends per ml.} \quad [3]$$

Values for b have been estimated from sedimentation data and from the behavior of λ molecules in solution: $b \approx 7.2 \times 10^{-2}$ μm (24). This leads to the simplification

$$j = \frac{63.4}{(\text{kb})^{1/2}} \ \mu\text{g/ml}, \quad [4]$$

where kb is the length of a particular DNA fragment in kb. If ligation is carried out at a DNA concentration, i, which is less than j, formation of circles will be favored. If $i > j$, multimer formation will be favored. At a given concentration, i, the fraction of circles formed is predicted by the equation

$$\% \text{ circles} = \frac{j}{i+j} \times 100. \quad [5]$$

The model system described in Fig. 2 was used to test these theoretical predictions for a 45-kb DNA molecule, for which j is calculated to be 9.4 μg/ml. A qualitative assessment of the products of ligation by gel electrophoresis of a restriction digest (Fig. 3) is consistent with the calculated value of j. More quantitative assessment at the lower concentrations, made by counting junction-fragment plaques (Fig. 4), also agrees well with theoretical predictions. Thus for a 45-kb molecule, a concentration of 0.8 μg/ml results in ≈90% circular fragments. The concentration needed to achieve a given percentage of circular fragments changes inversely as the square root of the length of the DNA molecule. For a 180-kb molecule, therefore, the concentration would need to be reduced to 0.4 μg/ml.

It should be apparent that any noncircular fragments will defeat the cloning strategy. These can be dealt with either (*i*) by screening junction fragments obtained with a given probe to be sure that they are syntenic, using somatic cell hybrid DNAs or *in situ* hybridization, or (*ii*) by adding a second size-selection step after the ligation to separate monomers from multimers.

The approach diagrammed in Fig. 1 includes the insertion of a marker gene within the circles to identify junction fragments, allowing exclusion of other fragments from the library. This step is not absolutely essential for use of the circularization method; one could screen a library of *Eco*RI fragments generated from circularized DNA with two probes that reside at either end of a single *Eco*RI fragment located in the genome at the starting point. Plaques that hybridized with both probes would not represent junction fragments and would not be studied whereas those that hybridized with one probe but not the other would be those in which the initial DNA cleavage had occurred. The disadvantage of this nonselected approach is that it requires screening a large number of plaques since only a small fraction will represent junction pieces. This could, however, be expedited using the recombination-selection system of Seed (25).

The use of a marker gene to select junction fragments has been investigated with a suppressor tRNA gene, though other biological markers such as the λ *cos* site or antibiotic resistance genes can be considered. The efficiency of incorporation of *supF* into covalent circles is 70–80% in the presence of a 500-fold molar excess of the marker gene. In the scheme shown in Fig. 2, it may be preferable that the marker gene length be shorter than 300 base pairs so that it cannot self-circularize (23) since such self-ligation will result in loss of the marker gene from the desired reaction.

A variation on the *supF* approach that would increase the size of the junction clones would be to perform a partial (rather than total) *Eco*RI digest of these marked circles and then clone the fragments into cosmids in which either the cosmid vector or the host carries an amber-mutated antibiotic-resistance gene.

DNA fragments allowing physical selection of the junction prior to λ cloning could also be used. Such selectable markers would include a DNA fragment labeled on one strand with biotin (26), allowing purification of junction pieces using a biotin antibody or avidin column, and a *lac* operator DNA fragment, allowing purification of junction pieces by operator-repressor binding (27). An additional advantage of this physical selection approach is that the resulting λ library

could still be selected by the *in vivo* recombination method (25), thus reducing the labor involved.

It will be possible to a large extent to select the direction in which one wishes to travel from the initial probe. A particular fragment can only be detected if there are no *Eco*RI sites between the probe and the *Mbo* I site where this fragment is ligated to the marker gene. Therefore, if the probe is short and has an *Eco*RI site as one endpoint, any junction fragments obtained by screening with this probe will have the arrangement *Eco*RI site-new DNA-*Mbo*I site-marker gene-*Mbo*I site-adjacent DNA-probe-*Eco*RI site, which will yield a clone in which the new DNA arises from the *Eco*RI side of the probe (refer to Fig. 1).

The applications of this method to directional scanning of the genome are numerous, especially with a series of libraries, each of which represents a different sized step. The method could, for example, be applied to the major histocompatibility locus (8) and the immunoglobulin loci (9). One could start with a known probe and take a directionally selected hop to generate a new probe. If this appeared to be in an interesting region (based on cross-homology with other members of that complex locus), further mapping could be done. On the other hand, if the new probe did not appear to be in an interesting region, a further step could be taken.

Another use of this approach would be to move closer to a genetic disease locus once a restriction-fragment-length polymorphism (6) has been identified that is linked to that locus. In Huntington disease, for example, a probe has been identified which is 3–5 centimorgans away from the Huntington disease locus (28, 29). A series of directional steps using libraries such as we describe should allow one to come closer to the disease locus and eventually obtain the gene responsible for the disease, as well as produce probes that, because of their tighter linkage to the disease gene, are more useful for diagnosis. The chromosome-hopping method can also be used to cross chromosome deletions that are too small to be visible cytogenetically, such as occur in a number of the hereditary persistence of fetal hemoglobin syndromes (reviewed in ref. 30), or to cross the junction in a translocated chromosome.

In conclusion, we present here the basis of a new method for obtaining clones at a considerable distance from an initial cloned fragment, without the need for overlapping clones of DNA regions between the two sites. We demonstrate with a model system that circularization of large DNA fragments, upon which the proposed method depends, is practical. The generation of junction-fragment libraries by this technique appears to be an achievable goal with many potential applications.

We thank Dr. Robert Dunn for supplying the plasmid pRD69 and the strain CARD-15, Dr. E. S. P. Reddy for assistance in preparing the *supF* gene, Drs. David Ward and Dan DiMaio for helpful discussions, Marsha Newman for excellent technical assistance, and Ann M. Mulvey for preparing this manuscript. F.S.C. acknowledges support from the Cooley's Anemia Foundation, Inc., and the Charles E. Culpepper Foundation, and S.M.W. acknowledges support by National Institutes of Health Grants CA30938 and AM 33871.

1. Collins, J. & Hohn, B. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 4242–4246.
2. Lau, Y. & Kan, Y. W. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 5225–5229.
3. Grosveld, F. G., Dahl, H. M., deBoer, E. & Flavell, R. A. (1981) *Gene* **13,** 227–237.
4. Steinmetz, M., Minard, K., Horvath, S., McNicholas, J., Srelinger, J., Wake, C., Long, E., Mach, B. & Hood, L. (1982) *Nature (London)* **300,** 35–42.
5. Carroll, M. C., Campbell, R. D., Bentley, D. R. & Porter, R. R. (1984) *Nature (London)* **307,** 237–241.
6. Shows, J. B., Sakaguchi, A. Y. & Naylor, S. L. (1982) *Adv. Hum. Genet.* **12,** 341–452.
7. Harper, M. E. & Saunders, G. F. (1981) *Chromsoma (Berlin)* **83,** 431–439.
8. Steinmetz, M. & Hood, L. (1983) *Science* **222,** 727–733.
9. Tonegawa, S. (1983) *Nature (London)* **302,** 575–581.
10. Blin, N. & Stafford, D. W. (1976) *Nucleic Acids Res.* **3,** 2302–2308.
11. Ish-Horowicz, D. & Burke, J. F. (1981) *Nucleic Acids Res.* **9,** 2989–2998.
12. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
13. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
14. Dunn, R. J., Belagaje, R., Brown, E. L. & Khorana, H. G. (1981) *J. Biol. Chem.* **256,** 6109–6118.
15. Russel, R. L., Abelson, J. N., Landy, A., Gefter, M. L., Brenner, S. & Smith, J. D. (1970) *J. Mol. Biol.* **47,** 1–14.
16. Fangman, W. L. (1978) *Nucleic Acids Res.* **5,** 653–665.
17. Serwer, P. (1980) *Biochemistry* **19,** 3001–3004.
18. Serwer, P. (1981) *Anal. Biochem.* **112,** 351–356.
19. Schwartz, D. C., Saffran, W., Welsh, J., Haas, R., Goldenberg, M. & Cantor, C. R. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47,** 189–195.
20. Schwartz, D. C. & Cantor, C. R. (1984) *Cell* **37,** 67–75.
21. Jacobson, H. & Stockmayer, W. H. (1950) *J. Chem. Phys.* **18,** 1600–1606.
22. Wang, J. C. & Davidson, N. (1966) *J. Mol. Biol.* **19,** 469–482.
23. Dugaiczyk, A., Boyer, H. W. & Goodman, H. M. (1975) *J. Mol. Biol.* **96,** 171–184.
24. Hearst, J. & Stockmayer, W. H. (1962) *J. Chem. Phys.* **37,** 1425–1433.
25. Seed, B. (1983) *Nucleic Acids Res.* **11,** 2427–2445.
26. Langer, P. R., Waldrop, A. A. & Ward, D. C. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 6633–6637.
27. Potter, H. & Dressler, D. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 1084–1088.
28. Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., Ottina, K., Wallace, M. R., Sakaguchi, A. Y., Young, A. B., Shoulson, I., Bonilla, E. & Martin, J. B. (1983) *Nature (London)* **306,** 234–238.
29. Cantor, C. R. (1984) *Nature (London)* **308,** 404–405.
30. Collins, F. S. & Weissman, S. M. (1984) *Prog. Nucleic Acid Res. Mol. Biol.* **31,** 315–458.